

Integration of MSFC Usability Lab with Usability Testing

Yiwei Cheng¹

Texas A&M University, College Station, Texas, 77840

and

Sally Richardson²

NASA MSFC, Huntsville, Alabama, 35812

Nomenclature

<i>SHE</i>	=	Safety, Health, and Environmental
<i>MSFC</i>	=	Marshall Space Flight Center
<i>NSRS</i>	=	NASA Safety Reporting System
<i>POC</i>	=	Point of Contact
<i>SCRS</i>	=	Safety Concerns Reporting System
<i>OSHA</i>	=	Occupational Safety and Health Administration

I. Introduction

As part of the Stage Analysis Branch, human factors engineering plays an important role in relating humans to the systems of hardware and structure designs of the new launch vehicle. While many branches are involved in the technical aspects of creating a launch vehicle, human factors connects humans to the scientific systems with the goal of improving operational performance and safety while reducing operational error and damage to the hardware. Human factors engineers use physical and computerized models to visualize possible areas for improvements to ensure human accessibility to components requiring maintenance and that the necessary maintenance activities can be accomplished with minimal risks to human and hardware. Many methods of testing are used to fulfill this goal, such as physical mockups, computerized visualization, and usability testing. In this analysis, a usability test is conducted to test how usable a website is to users who are and are not familiar with it. The testing is performed using participants and Morae software to record and analyze the results. This analysis will be a preliminary test of the usability lab in preparation for use in new spacecraft programs, NASA Enterprise, or other NASA websites.

The usability lab project is divided into two parts: integration of the usability lab and a preliminary test of the usability lab.

II. Integration of Usability Lab

¹MSFC MSGR Intern, EV 82—Stage Analysis Branch, MSFC, Texas A&M University

² Human Factors Engineer, EV 82—Stage Analysis Branch, EV 82, MSFC

From a human factors and psychology perspective, the MSFC usability lab should be optimum in the physical surroundings and task assessments from the Morae usability software in preparation for the preliminary usability testing.

A. Physical Environment

For optimal performance from the participants, the test laboratory is arranged with two workstations, one for the participant and one for the observer. The workstations are enclosed so that the participant and observer cannot physically see each other. Each station has a docking station for the laptop, a monitor, webcam, and microphone.

B. Morae Software

In preparation for usability testing, task and survey questions were developed after researching on the effects of wording and types of questions are the most reflective of the participants' opinions and are not misleading. The task questions and survey were programmed into Morae Usability Software when delivering the usability test to the participants. The task questions are designed to be reflective of the type of information people use the website to search for, and the survey questions are worded so that they are not leading and accurately reflect participants' viewpoint.

The Morae usability software used to conduct the test includes three parts to highlight data needed for analyzing usability. Morae Recorder captures audio, video, on-screen activity, and mouse and keyboard inputs during research. Morae Observer enables researchers to watch participants' experience, make notes, and mark tasks in real time. Morae Manager is used to view and analyze recordings, calculate metrics, such as average mouse clicks, generate graphs of the metrics, and create highlight videos from recordings.

III. Usability Testing

The purpose of usability testing is to observe potential users interacting with the product to discover errors and areas of improvement. This usability analysis focuses on the MSFC Safety, Health, and Environment (SHE) website as a preliminary test of the usability lab. Usability testing tasks and survey questions focus on how participants respond to four areas:

Performance: How much time and how many steps are required for people to complete basic tasks?

Accuracy: How many mistakes do people make?

Recall: How much does the person remember afterwards or after periods of non-use?

Emotional Response: How does the person feel about the tasks completed? Is the person confident or stressed?

The combination of the four areas reflects the ease of usability of the website.

A. Procedure

In this analysis of the SHE website, the type of testing chosen is key task testing, which asks users to do a task and then assessing how well they did. Key tasks on the website are selected to test how user-friendly the website is. Eighteen participants were separated into two groups, each with a different set of seven tasks. Morae usability software delivers the tasks beginning with the purpose and instructions of the usability test. Participants click begin task and end task as they start and finish a task. The software presents the survey after the end task of the last task.

B. Task Questions for Group 1

Task questions were chosen to be reflective of the type of information potential users may use the website to find. The tasks are representative of different heading tabs from the website.

1. Find the laboratory safety checklist an employee should review prior to working in a lab environment.
2. Show the steps in requesting an evaluation for more ergonomic furniture.
3. Find the roles of Safety, Health, and Environmental (SHE) organization.
4. Show the steps in reporting a chemical spill that happened on the east side of 4203.
5. Find the SHE representative and points of contact for safety issues for your organization.
6. Find newsletter articles on healthy, fitness, and environmentally friendly ideas for the month of June.

7. You missed a safety meeting for your branch. Find the last safety training presentation from the meeting.

C. Hierarchy for Answers

The multiple steps in reaching the information asked is critical in determining the problem and how to improve the website. This hierarchy is the optimum number of steps needed to take to reach the end of the task.

Task 1: News and Information→ SHE checklists→ Employee Laboratory Safety Checklist

Task 2: Services→ Ergonomics→ Ergonomic Evaluation Request

Task 3: News and Information→ SHE Roles and Responsibilities

Task 4: Mishaps, Questions, or Concerns→ SCRS→ Continue to SHE Report Form MSFC Main hazard/concern

Task 5: Who to Call→ SHE POC→ (look for organization code)

Task 6: News and Information→ Environmental Newsletters→ June

Task 7: Training→ SHE meeting topics→ June

D. Results for Group 1

From the four areas of performance, accuracy, recall, emotional response perspective, each metric reveals accuracy of the set of seven tasks. The overall success distribution classifies how successful the participants completed the tasks and is categorized into completed with ease, completed with difficulty, and failed to complete. Completed with difficulty is defined as clicking on one or more wrong tab or link before reaching the correct one. Failed to complete include answers that were not correct even though the user believed they were. From Figure 1, tasks 1, 2, 6, and 7 were easiest in terms of completion (the lowest percentage of failed to complete) while tasks 3, 4, and 5 were the hardest. Tasks 3 and 5 were the tasks users deemed their answer as accurate but were not.

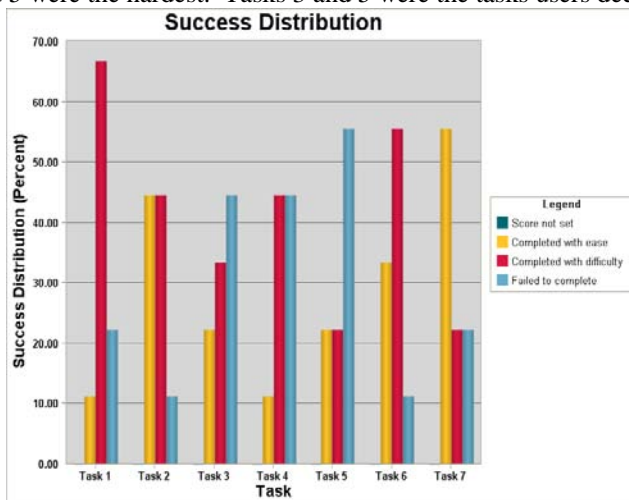


Figure 1. Success Distribution for Group 1

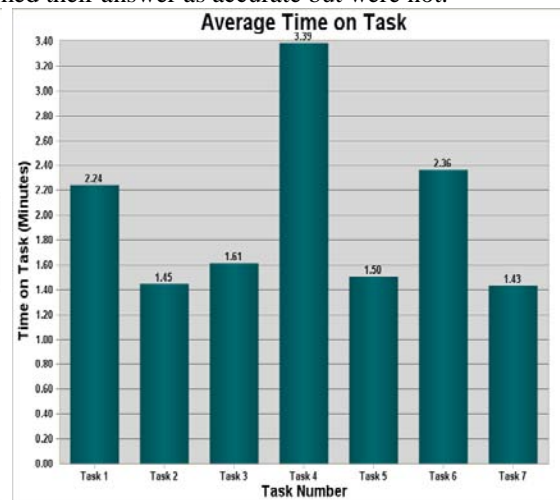


Figure 2. Average Time per Task for Group 1

The second metric used is the average time per task to measure performance and how long participants spent on each task. The highest average of time spent on task is task four. Task four may have taken the longest because participants looked around on website to find the link they are looking for. The main problem among the participants for task 4 is that participants were unsure of whether to go to NASA Safety Reporting System (NSRS) or Safety Concern Reporting System (SCRS). Task 3 and 5 were difficult for the participants because they were going to the wrong link under the main heading tab; however the wrong answers were consistent

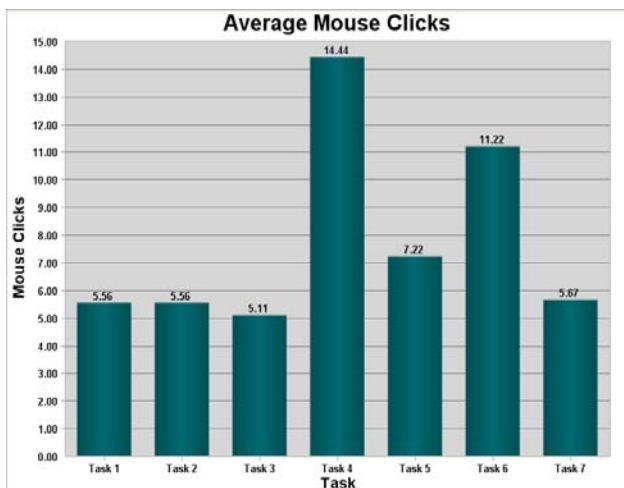


Figure 3. Average Mouse Clicks for Group 1

across the participants. Task 4 has the highest average time per task and the highest number of average mouse clicks. The number of mouse clicks varies from 2 to 41 times.

In analyzing the results, the level of hierarchy the participants were making the critical mistakes in is determined for the most difficult tasks. For task 3, participants had difficulty finding the main heading tab of News and Information. For task 4, all the participants went to the main heading of Mishaps, Questions, and Concerns, but the participants did not know whether to click SCRS or NSRS and which link to click from the drop down box. For task 5, most participants went under Who to Call tab but clicked the Building Manager Point of Contact instead of SHE POC.

E. Task Questions for Group 2

Task questions were chosen to be reflective of the type of information potential users may use the website to find. The tasks are representative of different heading tabs from the website.

1. Find the summary of work related injuries and illnesses documented each year.
2. Find articles on OSHA requirements and standards for the safety of asbestos.
3. Find a point of contact with access to a peroxide compound.
4. Report a hazard that happened days ago and was reported but inadequate action was taken to control it.
5. Find the lessons learned for the Space Shuttle at Marshall Space Flight Center.
6. Find all the accidents that happened in your building.
7. Find the list of training modules on occupational health.

F. Hierarchy for Answers

The multiple steps in reaching the information asked is critical in determining the problem and how to improve the website. This hierarchy is the optimum number of steps needed to take to reach the end of the task.

Task 1: Mishaps, Questions, or Concerns→ OSHA 300 Log

Task 2: News and Information→ OSHA Publications→ Asbestos→ (click Fact Sheet or any other related links)

Task 3: Databases and Links→ MSDS→ Chemical Inventory by Chemical→ (type peroxide)→ Search

Task 4: Mishaps, Questions, or Concerns→ NSRS→ How do I submit a report to NSRS? → (click .pdf)

Task 5:

Task 6: Mishaps, Questions, or Concerns→ Mishaps by Building Number→ (click on building number)

Task 7: SHE Organizations→ Environmental Engineering→ Training

G. Results for Group 2

From the group two results, tasks 1, 3, 4, and 7 were the hardest for the participants while 5 and 6 were the easiest. Users thought they had found the correct answer to tasks 1, 3, 4, and 7 but did not. Metrics used to

Standard deviation:	Number of Clicks
• 4.39—task 1	(2)
• 3.81—task 2	(2)
• 1.36—task 3	(1)
• 14.41—task 4	(7)
• 5.67—task 5	(1)
• 7.71—task 6	(2)
• 4.06—task 7	(2)

Figure 4: Standard Deviation of Mouse Clicks

determine ease of usability results are consistent throughout. Task 3 has the highest percent of incompleteness, second highest

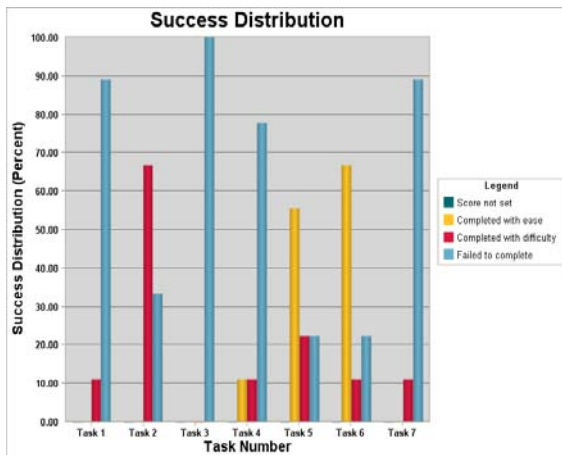


Figure 5: Success Distribution for Group 2

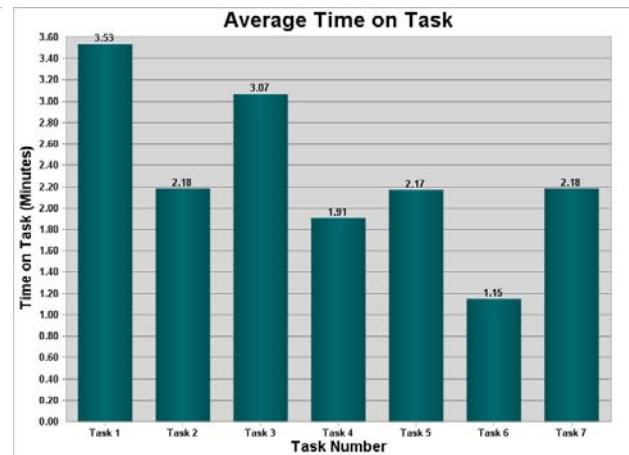


Figure 6: Average Time per Task for Group 2

average time spent on the task, and the highest number of average mouse clicks. Task 3 averaged fourteen mouse clicks with a standard deviation of about 13 clicks while the optimum number is four. Participants did not know what tabs to go to for finding a point of contact for the peroxide compound, and the wrong answers were not consistent. Task 1 also took a long time for the participants because they had to look around the website for tabs that may relate to summaries of work related injuries and illnesses. Some accidentally found the tab after clicking around the website for a while. Task 7 took a long time for the participants to find the occupational training modules because the information is placed in a different location than expected, so participants found similar training modules using a different procedure. Task 6 has the lowest average time because the tab name reflects the content and is in the expected logical location.

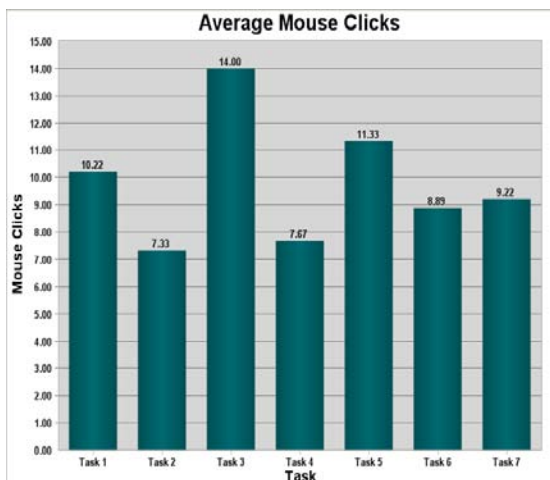


Figure 7: Average Mouse Clicks for Group 2

Standard deviation:	Number of Clicks
• 6.89—task 1	(1)
• 4.85—task 2	(3)
• 12.71—task 3	(4)
• 2.94—task 4	(3)
• 6.5—task 5	(6)
• 5.46—task 6	(3)
• 7.90—task 7	(2)

Figure 8: Standard Deviation of Mouse Clicks

In the level of hierarchy for task 1, many participants went under the Mishaps, Questions, and Concerns level, but did not understand OSHA 300 Log. From there, some found the webpage on accident and others clicked on another tab under Metrics and Statuses. Task 3's problem begins on the first level of hierarchy. The participants were not consistent in the process of going to the same wrong places. Each person has his own interpretation of where the information should be. Task 4 again reinforces the question of knowing the difference between NSRS and SCRS, and it is that level of his hierarchy in which most mistakes took place. Some participants for task 7 found the information on accident while others found the training module another way. That way led the

participants to basically the same training modules, except there were some additional ones to the training module website they should have went to.

H. Survey Questions

Each participant completes a survey at the end of the tasks. Even though participants have different set of tasks, the survey questions remain the same as they evaluate users' experience with the website.

1. How often do you visit the SHE website?
2. How easy is it for you to navigate the website for information?
3. All the tab names are informative.
4. I had to work hard to accomplish the tasks.
5. I was successful in accomplishing the tasks.
6. If you were not able to complete or had problems with any tasks today, why?
7. I was discouraged and/or annoyed while I was doing the tasks.
8. List 3 characteristics of the website that made it easy and fast to navigate.
9. List 3 characteristics of the website that made it difficult to navigate.
10. What aspects of the website should be changed to better help you find the information you need?

I. Survey Results

The analysis is grouped into two categories of frequently-occasionally and rarely-never.

1. How often do you visit the SHE website?
 0% Daily (A1)
 11% Frequently (A2)
 5.5% Occasionally (A3)
 33% Rarely (A4)
 50% Never (A5)

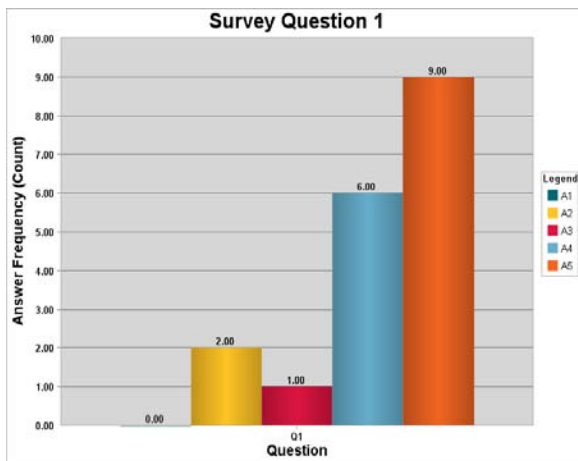


Figure 9. Survey Question 1

2. How easy is it for you to navigate the website for information?
 0% 1 (Easy) (A1)
 28% 2 (A2)
 33% 3 (A3)
 33% 4 (A4)
 5.5% 5 (Difficult) (A5)

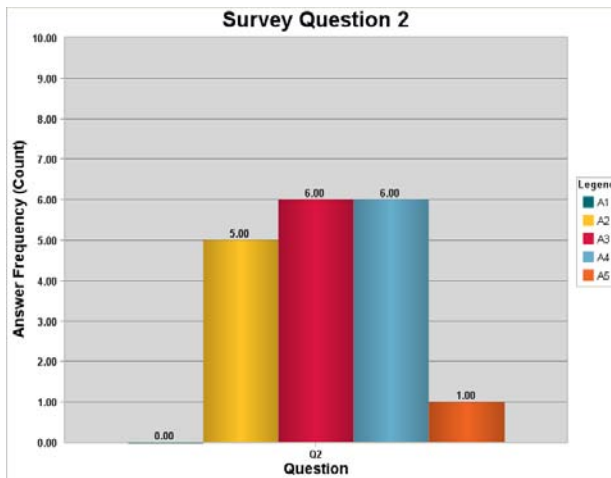


Figure 10. Survey Question 2

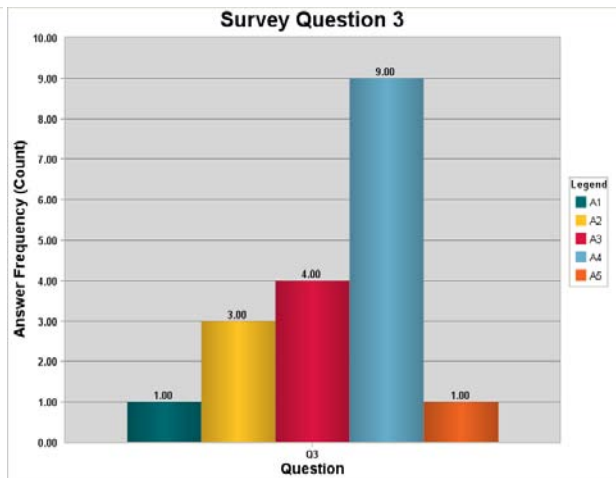


Figure 11. Survey Question 3

3. All the tab names are informative.
 - 5.5% Strongly Disagree (A1)
 - 17% Disagree (A2)
 - 22% Neutral (A3)
 - 50% Agree (A4)
 - 5.5% Strongly Agree (A5)
4. I had to work hard to accomplish the tasks.
 - 5.5% Strongly Disagree (A1)
 - 28% Disagree (A2)
 - 17% Neutral (A3)
 - 39% Agree (A4)
 - 11% Strongly Agree (A5)
5. I was successful in accomplishing the tasks.
 - 5.5% Strongly Disagree (A1)
 - 17% Disagree (A2)
 - 22% Neutral (A3)
 - 39% Agree (A4)
 - 17% Strongly Agree (A5)

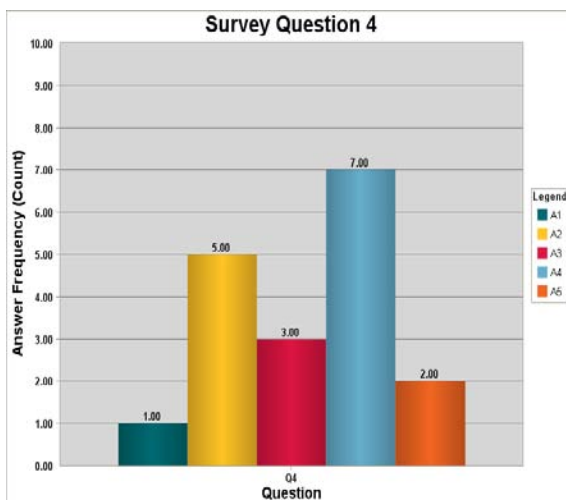


Figure 12. Survey Question 4

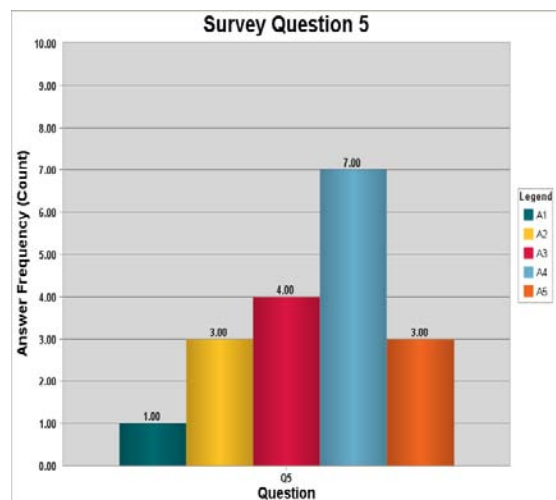


Figure 13. Survey Question 5

6. If you were not able to complete or had problems with any tasks today, why?

Frequently-Occasionally

- “I completed all of them.” (group 1)
- “No search box on SHE page. Tabs on SHE page not very descriptive.” (2)

Rarely-Never

- “there wasn't headings that matched as closely as I would have liked to the topic I was searching for.” (2)
- “The things I was looking for [weren't] listed as a page name under the tabs.” (2)
- “At times I felt like I was getting stuck in dead ends in terms of which links were appropriate, when it's possible I was on the right track in the first place.” (2)

7. I was discouraged and/or annoyed while I was doing the tasks.

5.5% Strongly Disagree (A1)
 22% Disagree (A2)
 28% Neutral (A3)
 39% Agree (A4)
 5.5% Strongly Agree (A5)

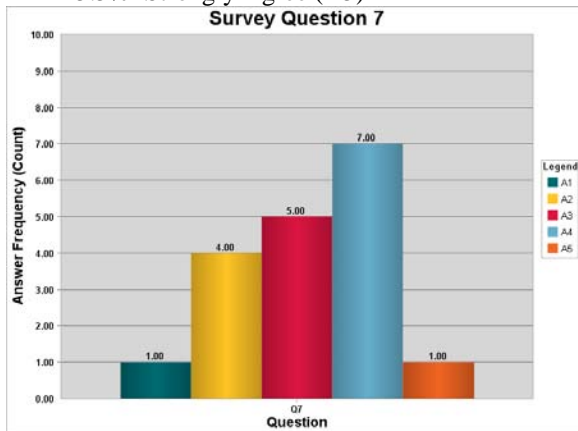


Figure 14. Survey Question 7

8. List 3 characteristics of the website that made it easy and fast to navigate.

Frequently-Occasionally

- “Clear headers are helpful Training is right up front Major tabs are fairly self explanatory” (1)
- “none!” (2)

Rarely-Never

- “Easy to understand language (don't have to have a lot of prior knowledge to understand what they were going for)
- Tabs were specific enough in description to get me to the right area to look
- The reporting button had a horn symbol by it that made it easy to pick out from other buttons.” (1)
- The pull down menus made searching the participant matter quick.” (1)

9. List 3 characteristics of the website that made it difficult to navigate.

Frequently-Occasionally

- “Too many items under certain topics
- Policies and Goals are in varying locations
- Sub menu items are sometimes documents, so those could almost be in a document list.” (1)

- “1. No search box. 2. The casual user does not care what bucket the information is in, we just want to find it quickly-webpage does not help do that. 3. Designed for 'expert users' not casual users.” (2)

Rarely-Never

- “Poor topic labeling, poor organization, apparent overlapping of information location. Also the fact that this is not a website that I frequent made it more difficult for me. (Unfamiliarity) (Information seemed to be tucked away in remote locations.)” (1)
- “I think that some of the links under the tabs were misplaced because they would have been suited better under some other tab. For example, not all of the SHE organizations was under the SHE tab” (2)

10. What aspects of the website should be changed to better help you find the information you need?

Top 5 Recommendations:

1. “More specific tab titles—may need more branches” (1)
2. “Search bar” (1)
3. “Fewer options to find information under each tab, the information given be less worded but detailed still.” (1)
4. “More standard layout with frequent links posted on left of page. Some menu items are in illogical locations.” (1)
5. “Show the directory path on each page.” (2)

Other Suggestions:

- “pop-up blurbs when you hover over different links”
- “The center was ignored; I focused on the Tabs in my search. I think it would be best to put the information listed in the middle screen in the tabs.”

IV. Conclusion

From the usability test, there are some recommendations to improve the usability of the website. The premise of usability testing is that there should not be any unnecessary thinking while attempting the task or navigating the website. The first recommendation is to eliminate any unnecessary thinking for the website, or in other words, how to use the website. Some participants asked questions like whether the main heading tabs can be clicked on for the drop down links to appear or what do the names on the tabs actually mean in terms of content. The second recommendation is to use more self-explanatory tabs. Instead of OSHA 300 Log, the title could be Summary of Work Related Injuries and Illness or related to the meaning of OSHA 300 Log. Similar information should be consolidated because many places on the website have related information but are in different locations. The consolidating will downsize the number of unnecessary tabs which will compensate for the increasing the number of more specifically titled ones. Regular maintenance on the website should be scheduled to identify problems and quickly solve them so that users will not be hindered in finding information. For example, many links to other web pages are broken and some building numbers are not on the map for accidents by building. An addition of a main search bar for keywords and topics will dramatically shorten the amount of time otherwise used to look around the website. The most important recommendation is to consider the users' recommendations for improvements so that the website can better help the users.

Some variances on average time per task and average mouse clicks per task may be due to not having a universal start and end page for the test participants. Future usability testing may take this into consideration as well as eliminate problems in the website such as not having 4205 on the accidents by building.

The preliminary testing creates a path for future work for the usability lab to test ground crew displays, payload displays, NASA Enterprise Applications, MSFC Employee Applications, labeling specifications for astronauts, or human interface display programs for future spacecrafts. The usability lab will be a valuable resource for the human factors team in making critical decisions for solving usability problems.

Acknowledgments

I would like to thank my mentor Sally Richardson for guiding me through the process of usability testing, Mark Blasingame for fixing all the computer problems, David Reynolds for editing task questions and my papers, Don Krupp for reviewing the presentation and test results, EV 82, and all the participants in the usability test.